Cumulative AP Practice Test 4 Solutions

Page 800

AP4.1 e.
AP4.2 c.
AP4.3 e.
AP4.4 a.
AP4.5 b.
AP4.6 e.
AP4.7 d.
AP4.8 a.
AP4.9 e.
AP4.10 a.
AP4.11 b.
AP4.12 d.
AP4.13 e.
AP4.14 d.
AP4.15 b.
AP4.16 a.
AP4.17 d.
AP4.18 c.
AP4.19 a.
AP4.20 b.
AP4.21 e.
AP4.22 e.
AP4.23 b.
AP4.24 c.
AP4.25 d.
AP4.26 e.
AP4.27 c.
AP4.28 c.
AP4.29 b.
AP4.30 b.
AP4.31 d.
AP4.32 a.
AP4.33 e.
AP4.34 b.
AP4.35 a.
AP4.36 b.
AP4.37 d.
AP4.38 a.
AP4.39 d.
AP4.40 d.

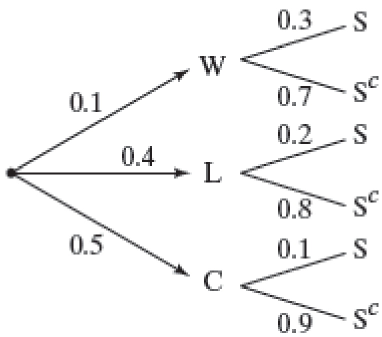
AP4.41 *State*: We want to perform a test of $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 \neq 0$ where μ_1 = true mean difference in electrical potential for diabetic mice and μ_2 = true mean difference in electrical potential for normal mice at the 5% significance level. *Plan*: If conditions are met, we should carry out a two-sample *t* test for $\mu_1 - \mu_2$. Random: The data come from independent random samples. 10%: n_1 = 24 is less than 10% of all diabetic mice and n_2 = 18 is less than 10% of all normal mice. Normal/Large Sample: The samples are both small (n_1 = 24 < 30 and n_2 = 18 < 30), but graphs of the data reveal no outliers or strong skewness so a two-sample *t* procedure is appropriate. *Do*: The test statistic is $t = \dfrac{(13.090 - 10.022) - 0}{\sqrt{\dfrac{4.839^2}{24} + \dfrac{2.915^2}{18}}}$ = 2.55. Using df = 24 – 1 = 23 and Table B, the *P*-value is between 2(0.005) = 0.01 and 2(0.01) = 0.02. *Using technology*: *t* = 2.55, df = 38.46, *P*-value = 0.0149. *Conclude*: Because the *P*-value of 0.0149 is less than α = 0.05, we reject H_0. There is convincing evidence that the true mean difference in electric potential for diabetic mice is different than for normal mice.

AP4.42 (a) We want to test $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 < 0$ where p_1 = the true proportion of women like the ones in the study who were physically active as teens that would suffer a cognitive decline and p_2 = the true proportion of women like the ones in the study who were not physically active as teens that would suffer a cognitive decline.
(b) A two-sample *z* test for $p_1 - p_2$.
(c) No. Because the participants were mostly white women from only four states, the findings may not be generalizable to women in other racial and ethnic groups or who live in other states.
(d) Two variables are confounded when their effects on the response variable (measure of cognitive decline) cannot be distinguished from one another. For example, women who were physically active as teens might have also done other things differently as well, such as eating a healthier diet. We would be unable to determine if it was their physically active youth or their healthier diet that slowed their level of cognitive decline.

AP4.43 (a) Because the first question called it a "fat tax", people may have reacted negatively because they believe this is a tax on those who are overweight. The second question provides extra information that gets people thinking about the obesity problem in the U.S. and the increased health care that could be provided as a benefit with the tax money, which might make them respond more positively to the proposed tax. The question should be worded in a more straightforward manner. For example, "Would you support or oppose a tax on non-diet sugared soda?"
(b) This method samples only people at fast-food restaurants. They may go to these restaurants because they like the sugary drinks and wouldn't want to pay a tax on their favorite beverages. Thus, it is likely that the proportion of those who would oppose such a tax will be overestimated with this method. A random sample of all New York State residents should be taken to provide a better estimate of the level of support for such a tax.

(c) Use a stratified random sampling method in which each state is a stratum.

AP 4.44 Let W = Mr. Worcester arrives first, L = Mr. Legacy arrives first, C = Dr. Currier arrives first, and S = the coffee is strong. The tree diagram below organizes the given information.



(a) $P(S) = (0.1)(0.3) + (0.4)(0.2) + (0.5)(0.1) = 0.16.$ There is a 0.16 probability that the coffee will be strong on a randomly selected morning.

(b) $P(C \mid S) = \dfrac{P(C \cap S)}{P(S)} = \dfrac{(0.5)(0.1)}{(0.1)(0.3)+(0.4)(0.2)+(0.5)(0.1)} = \dfrac{0.05}{0.16} = 0.3125.$ Given that the coffee is strong, there is a 0.3125 probability that it was brewed by Dr. Currier.

AP4.45 (a) A linear model is not appropriate. The scatterplot exhibits a strong curved pattern.
(b) Model B is better because the scatterplot shows a much more linear pattern and its residual plot shows no leftover patterns. The scatterplot for Model A still has a curved pattern and the residual plot has a leftover U-shaped pattern.
(c) $\widehat{\ln(\text{weight})} = 15.491 - 1.5222\ln(3700) = 2.984,$ thus $\widehat{\text{weight}} = e^{2.984} = 19.77\text{mg}.$
(d) About 86.3% of the variation in ln(seed weight) is accounted for by the linear model relating ln(seed weight) to ln(seed count).

AP4.46 (a) Let X = diameter of a randomly selected lid. Because X follows a Normal distribution, the sampling distribution of $\bar{x}$ also follows a Normal distribution. The mean of the sampling distribution is $\mu_{\bar{x}} = \mu = 4$ inches. Because 25 is less than 10% of all lids produced that hour, the standard deviation of the sampling distribution is $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{0.02}{\sqrt{25}} = 0.004$ inches.

(b) **Step 1: State the distribution and values of interest.** We want to find $P(\bar{x} < 3.99$ or $\bar{x} > 4.01)$ using the $N(4, 0.004)$ distribution. **Step 2: Perform calculations. Show your work.** The standardized scores for the boundary values are $z = \dfrac{3.99-4}{0.004} = -2.50$ and $z = \dfrac{4.01-4}{0.004} = 2.50.$ From Table A, the proportion of z-scores below $z = -2.50$ is 0.0062 and the proportion of z-scores above 2.50 is 0.0062. Thus, the proportion of z-scores less than –2.50 or greater than 2.50 is 0.0062 + 0.0062 = 0.0124. *Using technology:* 1 – normalcdf(lower: 3.99, upper: 4.01, μ: 4, σ: 0.004) gives an area of 0.0124. **Step 3: Answer the question.** Assuming that the machine is working properly, there is a 0.0124 probability that the mean diameter of a sample of 25 lids is less than 3.99 inches or greater than 4.01 inches.

(c) **Step 1: State the distribution and values of interest.** We want to find $P(4 < \bar{x} < 4.01)$ using the $N(4, 0.004)$ distribution. **Step 2: Perform calculations. Show your work.** The standardized scores for the boundary values are $z = \frac{4-4}{0.004} = 0$ and $z = \frac{4.01-4}{0.004} = 2.50$. From Table A, the proportion of z-scores below $z = 0$ is 0.5000 and the proportion of z-scores below 2.50 is 0.9938. Thus, the proportion of z-scores between 0 and 2.50 is 0.9938 – 0.5000 = 0.4938. *Using technology:* normalcdf(lower: 4, upper: 4.01, μ: 4, σ: 0.004) gives an area of 0.4938. **Step 3: Answer the question.** Assuming that the machine is working properly, there is a 0.4938 probability that the mean diameter of a sample of 25 lids is between 4.00 and 4.01 inches.

(d) Let Y = the number of samples (out of 5) in which the sample mean is between 4.00 and 4.01. The random variable Y has a binomial distribution with $n = 5$ and $p = 0.4938$. $P(Y \geq 4) = P(Y = 4) + P(Y = 5) = \binom{5}{4}(0.4938)^4(1-0.4938)^1 + \binom{5}{5}(0.4938)^5(1-0.4938)^0 = 0.1798$. *Using technology*: $P(X \geq 4) = 1 - P(X \leq 3) = 1 -$ binomcdf(trials: 5, p: 0.4938, x value: 3) = 0.1798.

(e) Because the probability found in part (b) is less than the probability found in part (d), getting a sample mean below 3.99 or above 4.01 is more convincing evidence that the machine needs to be shut down. This event is much less likely to happen by chance when the machine is working correctly.

(f) Answers will vary. For example, stop the production process if there are 5 consecutive sample means between 4.00 and 4.01. Using this rule, the probability the machine will be shut down when working correctly is $(0.4938)^5 = 0.029$. Another possibility is to consider getting two consecutive samples means between 4.005 and 4.01. Note that $P(4.005 < \bar{x} < 4.01) = 0.0994$. If two consecutive samples are taken, then

$$P(\text{both between 4.005 and 4.01}) = (0.0994)^2 = 0.0099 \approx 0.01.$$